

Combining robustness and recovery in rapid transit network design

L. Cadarso^a and Á. Marín^b

^aDepartamental III Building, Rey Juan Carlos University, Fuenlabrada, Spain; ^bTechnical University of Madrid, Escuela Técnica Superior de Ingeniería Aeronáutica y del Espacio, Madrid, Spain

ABSTRACT

When designing a transport network, decisions are made according to an expected value for network state variables, such as infrastructure and vehicle conditions, which are uncertain at a planning horizon of up to decades. Because disruptions, such as infrastructure breakdowns, will arise and affect the network on the day of operations, actions must be taken from the network design. Robust network designs may be implemented but they are extremely expensive if disruptions do not realise. In this paper, we propose a novel approach to the network design problem where robustness and recovery are combined. We look for the trade-off between efficiency and robustness accounting for the possibility of recovering from disruptions: recoverable robust network design. Computational experiments drawn from fictitious and realistic networks show how the presented approach reduces the price of robustness and recovery costs as compared to traditional robust and non-robust rapid transit network design approaches.

ARTICLE HISTORY

Received 30 January 2015

Accepted 17 November 2015

KEYWORDS

Network design; disruption; recoverable robustness; rapid transit

1. Introduction

Designing a rapid transit network (RTN) or even extending one that is already functioning, is key to provide high quality of life to the population due to the fact that it reduces congestion, travel time and pollution. A RTN design (RTND) is highly dependent on the future system usage. According to Vuchic (2005), maximising passenger attraction is the most appropriate objective to consider when planning transit systems. Also, Gutiérrez-Jarpa et al. (2013) state that a frequently used objective in RTND is to maximise the covered population. Maximising population coverage requires relatively few data and usually yields tractable models (Curtin and Biba 2011; Escudero and Muñoz 2009; Laporte et al. 2011; Marín 2007; Marín and García-Ródenas 2009). However, the network design must be also as cheap as possible (i.e. construction costs). When designing a new network, the infrastructure designer must account for the fact that passengers will commute within it if the travel time is shorter than within the current options: when facing a RTND problem there is usually another transportation system already operating in the area where the RTN is to

be built or extended. The multimodal aspect of the problem is crucial in network design; therefore, the current transportation system must be taken into account.

The RTND problem aims at maximising the demand coverage by the new network subject to design and budget constraints, all while considering passengers' decisions when evaluating different travelling alternatives. Roughly speaking, the set of possible passenger decisions will be composed of two elements: first, travelling within the old (current) network, and second within the new (to be constructed) network. The comparison of these alternatives is made by passengers who evaluate and compare the generalised cost in the current and new networks.

Transportation systems are critical infrastructure systems, whose smooth operation is important for maintaining normal functions of our society. These spatially distributed systems are vulnerable to disruptions: during daily operations within a RTN different incidents may occur. Despite the unpredictable nature of them in terms of location, time, and magnitude, effective mitigation methods from an engineering perspective should be designed.

The first task after noticing an incident is to determine whether or not it requires a substantial active intervention. If it does, the operations are said to be disrupted, and recovery plans must be designed in order to minimise negative disruptive effects and recover operations from the disrupted situation as soon as possible (i.e. disruption management). Recently, a significant amount of research has been made related to disruption management in network design and operations.

Szeto and Lo (2005) states that planning over a long horizon faces uncertainty. Dealing with these uncertainties is a key factor for providing a resilient network for daily operations. The design of the network is determinant in order to absorb disruptive negative effects such as the loss of passengers due to low performance during the disruption. The more the network is able to absorb these disruptive negative effects, the more resilient the network is. In order to find resilient network designs, different research techniques such as stochastic programming (two-stage) and robust optimisation may be applied (Liebchen et al. 2009):

- A two-stage stochastic programme is a programme where part of the input data is subject to uncertainty. This uncertainty is characterised by a distribution which is either known, or partly known, or can at least be sampled. The decision variables split into first-stage decisions and second-stage decisions. The first-stage decisions must be chosen fixed for all scenarios (i.e. one scenario for each disruption). The second-stage variables can be chosen after the actual value of the random variables is revealed. Usually, the objective function of a two-stage stochastic programme comprises a cost function for the first-stage variables and the expected cost of the second-stage variables according to the given distribution.
- Robust optimisation considers a quite similar situation; however, there are not second-stage actions. Again, a certain part of the data is subject to uncertainty. But as the robust programme features no second-stage variables, the variables (which are fixed before the actual data are revealed) must form a feasible solution for every scenario. A robust solution fits for all scenarios. Likewise, the objective function of a robust programme contains no expectation or other stochastic component. The objective is a deterministic function of the solution. In robust modelling, it suffices to know the range in which the uncertain data can fluctuate. Usually, extremely unlikely scenarios are excluded and this uncertainty range is smaller than given in reality.

The approaches we have presented in order to find resilient networks (i.e. stochastic programming and robust optimisation) have major drawbacks: first, stochastic programming requires extensive knowledge of the probability distribution of disturbances; and second, robust optimisation requires a solution which has to be feasible for a large set of possible modifications of the original input (i.e. it might be far too conservative, and it does not account for available recovery actions).

The idea of providing solutions which are immune to data uncertainty is contained in robustness concepts. However, the required immunity may be less strict than the one formalised in robustness concepts: solutions may be recovered after data perturbation. The key is that the recovery of the system may not be so expensive as the introduction of traditional robustness concepts. Then, a new concept considers the integration of both, robustness and recovery. This new concept is the recoverable robustness: it studies the robustness of the system accounting for recovery actions for disrupted scenarios.

Recoverable robustness appears to overcome the above presented drawbacks: it explicitly takes into account available recovery actions and does not need to know probability distributions. Now, the solution of the problem does not need to be feasible for a whole set of admissible disturbances anymore, but it must be able to be recovered (with a limited cost) from a set of disrupted scenarios to get a feasible solution. This concept fits very well for network design problems: in practice, the probability distribution of disruptions is not known, so stochastic programming cannot be applied. Furthermore, traditional robust modelling lead to solutions that are far too conservative and too expensive. Recoverable robustness seems to be a natural approach to network design problems: design the network in such a way that the disruptive negative effects in the operational phase are limited when applying a given recovery strategy. Here strategy refers to the set of available actions during the network design phase (i.e. construction of infrastructure), years in advance to the day of operations.

In this paper, we present a new recoverable robust approach to RTND problems. We use the framework proposed in Laporte et al. (2007), where the core network is to be decided first and secondary stations are to be decided later (i.e. once the network topology is known, the secondary stations are located). In our approach, recoverable robustness means that a new network is to be designed maximising the new network trip coverage in the presence of a set of disruptions, which is limited, while considering recovery strategies with a limited impact. The word 'limited' has two different meanings here. First, there are disruptions of different nature (e.g. earthquakes vs. system malfunction); we focus on disruptions which only affect the network to be designed such as infrastructure malfunction, vehicle breakdowns and collisions. Second, the problem we study is at the strategic level and new infrastructure cannot be constructed in order to alleviate the negative disruptive effects once a disruption has started; therefore, the recovery has a limited impact in the network design problem. The rest of the impacts, which are related to the day of operations through a change in operation conditions (e.g. different passenger flows may require different capacities), are out of the scope of this paper.

This paper is organised as follows. In Section 2 the literature overview is given. Section 3 describes the traditional RTND problem and mathematical model in detail. Section 4 presents the traditional robust RTND (RRTND) model. Section 5 is devoted to the recoverable RRTND mathematical model. In Section 6 we present our computational experiments. Finally, we draw some conclusions.

2. Literature review

The optimisation of network design problems has been a fruitful area of applied operations research in the past decades. Numerous papers address and solve theoretical and real-life network design problems. Under the title of RTND, bus, railway or metro network designs are addressed. Several studies have addressed Bus RTN problems, generating routings but without including the network topology design. Wei and Machemehl (2004) and related works modelise a bus transit network where the aim is to minimise user trip time and exploitation costs and maximise the coverage of the demand with no more than two-transfers, considering the capacity of the resources. Leiva et al. (2010) deal with a Bus Rapid Transit system where the aim is to know which type of services should be offered, at what frequencies and with which type of vehicles. The model is formulated as a nonlinear programming problem where vehicle capacities are relaxed. Mauttone and Urquhart (2009) extend the previous works to determine the optimal traces of the routes. The generation of the routes and their frequencies is modelled as a multi-objective combinatorial optimisation problem. Walteros, Medaglia, and Riaño (2013) combine the meta-heuristic and exact approaches for generating routes and frequencies that minimise operative and passenger costs under system constraints.

Some other research address and solve theoretical and real-life railway network design problems. In the context of rail RTND, a complete review is given by Laporte and Mesa (2014). The first applications concentrated on solving simple network design problems composed of one single line and considering static and deterministic problem setting. Bruno, Gendreau, and Laporte (2002) are among the first authors in studying RTND problems; they present a mathematical model and a two-phase heuristic for the location of a rapid transit alignment in an urban setting; they minimise the user cost while the coverage of the demand by the new network is maximised. Laporte et al. (2007) extend the previous research: they incorporate the station location and the design of multiple lines; they define the model by maximising public demand coverage subject to budget constraints as side constraints. Marín (2007) studies the inclusion of free but bounded number of lines; here each line's origin–destination is freely chosen by the model.

Robustness concepts have been proposed by different authors. Bertsimas and Sim (2003) propose an approach to control the level of conservatism in the solution. Ben-Tal and Nemirovski (1999), Ben-Tal and Nemirovski (2002) and Bertsimas and Sim (2004) use the ellipsoidal uncertainty set concept; they consider the worst-case. However, these solutions are too conservative (i.e. they are expensive on a daily basis). Other robustness concepts have been proposed by Fischetti, Salvagin, and Zanette (2011), García et al. (2007) and Laporte et al. (2011). Li, Li, and Lam (2014) propose a model for integrated design of a sustainable land use and transportation system with uncertainty in future population. The future population in the urban area is assumed to be a random variable with a given probability distribution. The model is formulated as a two-stage robust optimisation problem. The first stage is to optimise the land use and transportation system by maximising a robust risk-averse objective function subject to various chance constraints. The second stage, after the future population has been determined, is a scenario-based stochastic location and route choice equilibrium problem. Szeto and Wang (2015) propose reliability-based system-optimal traffic assignment under supply uncertainty based on the concept of the total system travel time

budget, and defines the price of anarchy for the corresponding user equilibrium traffic assignment.

Nielsen, Kroon, and Maróti (2012), Cadarso, Marín, and Maróti (2013) and Cadarso, Maróti, and Marín (2015) study the railway recovery problem once a disruption has started. Nielsen, Kroon, and Maróti (2012) describe a generic framework for dealing with disruptions. This framework is presented as an online combinatorial decision problem, where the uncertainty of a disruption is modelled by a sequence of information updates. Cadarso, Marín, and Maróti (2013) present an integrated optimisation model for rescheduling rolling stock plans while accounting for passenger demand use. This model is extended in Cadarso, Maróti, and Marín (2015) in order to obtain schedules which are controllable and easier to be implemented by the operator.

Liebchen et al. (2009) develop the recoverable robustness concept within the railway industry; their main focus is on finding recoverable solutions with a limited effort. In case of disruption, they apply a recovery algorithm while minimising the maximum deviation of the recovered solutions from the planned solution. The deviation is a measure of the effort required to modify the planned solution (but affected by the disruption) into a recovered solution. They use probabilities for the occurrence of different disruptions.

Cicerone et al. (2009) study recoverable robust approaches for several railway shunting and timetabling problems. They analyse some recovery algorithms and provide lower and upper bounds to the solution. Caprara et al. (2008) propose an exact method for computing recoverable robust solutions for the train platforming problem by optimally buffering a network. Cacchiani et al. (2012) describe a two-stage optimisation model for determining recoverable robust rolling stock circulations for passenger trains. They use a Benders-based heuristic so as to solve the problem; first using Benders decomposition to determine the optimal solution for the linear relaxation of the model; and second, using the generated cuts to compute heuristic solutions for the mixed-integer programming model. Cadarso and Marín (2012b) study recoverable robustness for a RTND model. They consider several disrupted scenarios with their associated probabilities and minimise the planned and recovery costs. Planned costs are defined as passengers lost to the current network, routing costs and location costs in the undisrupted scenario. Recovery costs are a weighted sum of the passengers lost to the current network, routing costs and deviations in passenger flows within the new network with respect to the undisrupted scenario. Here, the weights are the probabilities of the scenarios. The authors show for small-sized case studies that improved network designs can be obtained according to the given probabilities. Codina, Marín, and Cadarso (2014) consider a network design approach based on failure probabilities. These are considered to be function of the level of service and/or rolling stock routing through the network. However, neither level of service nor rolling stock operation can be known beforehand, they result from the design process itself. A two recourse stochastic programming model is formulated where the failure probabilities are implicit functions of the number of services and rolling stock routing of the transit lines.

In the airline industry, schedule disruptions are also frequent causing negative disruptive effects such as flight cancellations and delays. Recoverable robustness has been also addressed within this topic. Froyland, Mahen, and Wu (2013) formulate a recoverable robust aircraft tail assignment model; they solve the model using a two-phase algorithm within Benders decomposition and incorporating enhancement techniques; they compare their algorithm with the method used by Grönkvist (2005), demonstrating that they can reduce

recovery costs. Eggenberg (2009) states that better schedules are less sensitive to perturbations and are easier to recover when severe disruptions occur; he presents a general model that encodes all feasible recovery schemes (i.e. a new route, pairing or itinerary) of any unit (i.e. an aircraft, a crew member or a passenger) of the recovery problem.

2.1. Contributions

In this paper, we study the recoverable RRTND problem. The network to be designed will maximise passenger trip coverage in the presence of a set of admissible disruptions while considering recovery.

The approach in this paper is based on the recoverable robust network design model proposed by Cadarso and Marín (2012b). They assume that the recoverable robust design is defined considering several disrupted scenarios with their associated probabilities. However, it is often very hard to know the probabilities, unless statistical data on past disruptions are available. Moreover, data on past events are very difficult to be used for the development of new infrastructure. Compared to that paper, the novelty of the current paper lies in the following aspects.

- We present a new approach which introduces the recoverable robustness concept in the field of RTND problems while (1) considering a limited set of disrupted scenarios with unknown probabilities of occurrence and (2) assuming that recovery, which has a limited impact on the system, is available.
- We develop a new formulation whose aim is to provide robust solutions which can be also be recovered with a limited effort; consequently, we address the minimisation of the maximum deviation (i.e. the effort that needs to be made so as to recover).
- We carry out computational experiments on fictitious and realistic instances and show how the price of robustness and recovery costs are reduced as compared to traditional robust and non-RRTND approaches.
- We develop a multiobjective study considering the trade-off between efficient network design (minimising passenger and infrastructure manager costs) and recoverable robust network design.

We acknowledge that the design of public transportation in a city is done in an integrated way. That is, the design must account for the new RTN and the new configuration of the current network. For example, being the current network a bus network, it will surely change when developing a new metro network. However, when we have the opposite case, where the current network is a metro network and the new network to be developed or extended is a bus network, it is not easy to change the current network. Nevertheless, this is not a drawback of our approach. If the current network is to be changed, a ‘what if’ approach may be developed. That is, given different current networks (i.e. expected current networks), see which is the optimal new network.

3. The RTND problem

The aim of the RTND problem is to design a new network, that is, to decide at which nodes to locate the stations and how to connect them, in order to cover as many passenger

trips as possible while minimising routing and construction costs. Recall that the current transportation network (e.g. a bus network) plays an important role: passengers will evaluate the different networks and will commute within the most convenient network.

The rest of this section introduces the network infrastructure, the passenger demand and the traditional mathematical formulation for a network design problem.

3.1. The network infrastructure

The RTN consists of arcs and nodes. We have two different types of nodes denoted by N : centroids (N_c) and potential stations (N_r). Centroid nodes are those where the demand is generated or attracted to. Potential station nodes are those where the network is built on and where passenger demand enters and leaves the network. We assume that the location of the potential stations is given. Then, we model the infrastructure as a graph with nodes $i \in N$, and with the set A of arcs linking them. Arcs may represent: potential alignments in the RTN (A_r), dummy arcs between origin centroids and any station (A_o), dummy arcs between stations and every destination centroid (A_d), and arcs between any origin–destination pair, corresponding to current network (A_f). The set of arcs is then $A = A_r \cup A_f \cup A_o \cup A_d$.

Therefore, we have a potential network (N, A) from which the optimum RTN is selected. Figure 1 shows an example of a potential network of 14 nodes: 4 potential stations and 10 centroids. The different types of arcs are represented. Each potential station i has an associated construction cost m_i and each potential arc $ij \in A_r$ a pair of weights (c_{ij}, d_{ij}) : the construction cost c_{ij} and the distance d_{ij} . The arcs are used to define the new lines to be constructed: L is the set of new lines. The overall planning problem for a transport network is usually sequentially solved (Cadarsó and Marín 2012a; Michaelis and Schöbel 2009). The paper presented here deals with the network design problem. But it also accounts for some details of the line planning problem. However, the line planning problem is about deciding origins and destinations of the lines and the frequency to be operated on them (i.e. capacity). Because we do not care about frequencies we cannot say we solve the line planning problem and therefore we are studying an uncapacitated network design problem. Then,

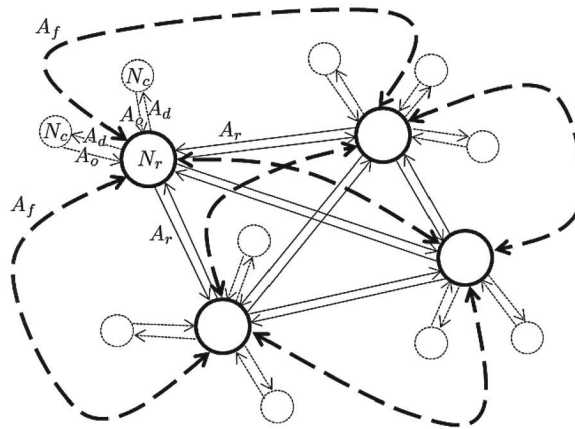


Figure 1. Example of a potential network.

we say the lines are elements to support the network design; elements to be accounted for when solving the line planning problem. The set of new lines L builds up the new network. The information known is the cardinality of the set, that is, an upper bound to the number of new lines to be constructed. However, this does not mean we give the description of the lines (arcs belonging to a line), this will be decided within the mathematical model. For a comprehensive study of this characteristic of the model, see Marín (2007). Since resources are limited a budget constraint on construction costs is also imposed.

As we are studying a real-life problem, the new infrastructure to be constructed will not be isolated from the current network. We will consider the existence of a current transport network formed by different modes of transport (e.g. a bus network). This current network has some stations in common with the RTN to be constructed. However, the networks are independent and they use different infrastructure. Again, the current network is given by the arcs between any origin–destination pair in A_f .

3.2. Passenger demand

The demand is characterised by an origin and a destination. We assume that the mobility patterns in the metropolitan area under study are known. This implies that the number of potential passengers from each origin to each destination is in average given. We define passenger groups as follows: $w = (o(w); d(w); g_w)$, where $o_w \in N$ is the origin centroid, $d_w \in N$ the destination centroid and g_w is the passenger group size (i.e. number of passengers). We acknowledge that the demand is time-dependent. However, we study an uncapacitated network design problem where the topology of the network is to be decided and time-dependent aspects of the problem are out of scope (Cadarso and Marín 2015; Codina, Marín, and Cadarso 2014; Escudero and Muñoz 2009; García et al. 2007; Laporte et al. 2007, 2011; Laporte and Mesa 2014; Marín 2007; Marín and Jaramillo 2008; Marín and García-Ródenas 2009).

The demand will be realised through the available paths in the new network or through a path within the current network. Each passenger group $w \in W$ will choose a path. Each path is characterised by its origin, destination and the arcs belonging to it. The demand will choose its path based on the generalised travel cost. Here, the generalised cost is defined as the distance each passenger group must travel to reach its destination (i.e. it is a non-monetary cost of the trip). This distance is equal to the sum of all the arcs' distances (d_{ij}) in the path for the new network, and equal to u_{cur}^w for the current network (alternative modes of transport). The problem we are studying is strategic and generalised costs are not well known; congestion is an important issue that must be accounted for. Although the distance to be travelled by passengers is independent from congestion, travel time is highly dependent on it. Therefore, we introduce a congestion parameter μ_w for each passenger group w which accounts for congestion so as to adjust generalised costs. The effect of congestion is only considered in the current generalised cost, which corresponds to the current network. That is, it measures the expected level of congestion of the current network. Because this paper does not consider line frequency planning, it is reasonable to consider that non-congested networks are being designed. This is also consistent with the fact that we study rapid transit systems, where there is usually an isolated infrastructure for the transport system (e.g. metro and rapid bus systems). Therefore, generalised costs in the new network are assumed to be independent of the level of congestion. Cadarso and Marín (2015), Escudero

and Muñoz (2009) and Marín (2007) study the effects of different expected levels of congestion on similar problems. The values for these parameters are to be estimated by the infrastructure designer and depend on the future state of the current network.

We assume that passengers will choose the path with the lowest associated generalised cost. We acknowledge that this and the fact that we include neither passenger types nor social and environmental attributes in the generalised cost function is a simplistic way of modelling passenger behaviour. However, the purpose of this work is to study the impact of disruptions through the concept of recoverable robustness on the network design phase. Therefore, the state of the network for each disruption (i.e. the perceived utility of each of the alternatives) is highly uncertain: the frequency, the timetable and vehicle capacity are not known at the network design phase. Hence, we assume this simplistic demand modelling approach.

3.3. The RTND model

The RTND problem aims at computing a network design in order to decide at which nodes to locate the stations and how to connect them so as to attract as many passengers as possible to the new network and minimise routing and construction costs (Cadaro and Marín 2015; Escudero and Muñoz 2009; Laporte et al. 2007; Marín 2007). There are constraints related to the infrastructure design, the passenger demand and the allocation of the demand to the network to be constructed.

Before introducing the mathematical model, we need to define the following notations:

Sets

- N is the set of nodes indexed by i and j .
- A is the set of arcs linking the nodes. It is indexed by ij (its origin node i and its destination node j).
- N_c is the subset of centroids.
- N_r is the subset of potential stations.
- $N(i) = \{j : \exists ij \in A\}$ is the subset of nodes adjacent to node i .
- A_r is the subset of potential arcs linking the stations.
- A_o is the subset of dummy arcs between origin centroids and stations.
- A_d is the subset of dummy arcs between stations and destination centroids.
- A_f is the subset of arcs between any origin–destination in the current network.
- W is the set of passenger groups indexed by w . The origin and destination centroids for each passenger group w are given by $o(w)$ and $d(w)$, respectively.

Parameters

- d_{ij} is the distance of each arc ij . It usually corresponds to the Euclidean distance between the nodes if the system is underground and the street distance if it is at ground level. However, forbidden regions will increase the distance.
- g_w is the number of passengers in passenger group w .
- c_{ij} is the construction cost for potential arc ij .
- m_i is the construction cost for potential station i .
- c_{\max} is the available budget for construction.

- μ_w is the congestion factor for each passenger group w in the current network.
- u_{cur}^w is the generalised cost (i.e. distance) for passenger group w through the current network.
- α is the weight in the objective function for the number of passengers using the current network.
- β is the weight in the objective function for location costs.
- γ is the weight in the objective function for routing costs.

Variables

- $x_{ij}^l \in \{0, 1\}$ equals 1 if line l is located using the arc ij ; 0, otherwise.
- $\chi_{ij} \in \{0, 1\}$ equals 1 if arc ij is located; 0, otherwise.
- $\psi_i \in \{0, 1\}$ equals 1 if station i is located; 0, otherwise.
- $y_i^l \in \{0, 1\}$ equals 1 if line l is located using the node i ; 0, otherwise.
- $h_l \in \{0, 1\}$ equals 1 if line l is located; 0, otherwise.
- $f_{ij}^w \in \{0, 1\}$ equals 1 if passenger group w uses arc ij in the new network; 0, otherwise.
- $f_{cur}^w \in \{0, 1\}$ equals 1 if passenger group w uses the current network; 0, otherwise.

The objective function and the constraints follow.

3.3.1. Objective function.

The traditional RTND model formulation minimises the number of passengers using the current network, location costs and routing costs. Hence, we deal with a multiobjective model. As usual, we minimise a positive, linear combination of the different costs (Marín and Jaramillo 2008):

$$\min z = \alpha z_{cur} + \beta z_{loc} + \gamma z_{route}, \quad (1)$$

where

$$z_{cur} = \sum_{w \in W} g_w f_{cur}^w \quad (2)$$

$$z_{loc} = \sum_{ij \in A_r, i < j} c_{ij} \chi_{ij} + \sum_{i \in N_r} m_i \psi_i, \quad (3)$$

$$z_{route} = \sum_{w \in W} \left(\sum_{ij \in A_r \cup A_d \cup A_o} (d_{ij} f_{ij}^w) + \mu_w u_{cur}^w f_{cur}^w \right). \quad (4)$$

Passenger trip coverage is the main objective in this problem: Equation (2) accounts for it. The location cost in Equation (3) is also minimised in order to avoid the construction of inoperative parts of network. Finally, the sum of the demand routing costs per passenger group is minimised in Equation (4).

3.3.2. Infrastructure constraints.

$$z_{loc} \leq c_{max}. \quad (5)$$

Constraint (5) limits the total construction cost to the available budget.

$$\chi_{ij} = \chi_{ji} \quad \forall ij \in A_r : i < j, \quad (6)$$

$$x_{ij} \leq \psi_i \quad \forall i \in N_r, \forall ij \in A_r : i < j, \quad (7)$$

$$x_{ij} \leq \psi_j \quad \forall j \in N_r, \forall ij \in A_r : i < j. \quad (8)$$

Constraints (6)–(8) ensure that the location of nodes and arcs is coherent: for each located arc both riding directions are available, and every arc must have its origin and destination nodes located.

$$x_{ij}^l \leq y_i^l \quad \forall ij \in A_r : i < j, \forall l \in L, \quad (9)$$

$$x_{ij}^l \leq y_j^l \quad \forall ij \in A_r : i < j, \forall l \in L, \quad (10)$$

$$x_{ij}^l = x_{ji}^l \quad \forall ij \in A_r : i < j, \forall l \in L, \quad (11)$$

$$y_i^l \leq \psi_i \quad \forall i \in N_r, \forall l \in L, \quad (12)$$

$$\sum_{l \in L} x_{ij}^l = x_{ij} \quad \forall ij \in A_r : i < j. \quad (13)$$

Constraints (9)–(12) are arc-line location constraints. They are similar to constraints (6)–(8). Constraints (13) link the arc-line variables with the arc-location variables. In addition, line-path constraints may be imposed in order to guarantee that lines either enter and exit each station or that only enter/exit to/from it (see Marín 2007 for more details on this).

$$\sum_{ij \in A_r : i < j} x_{ij}^l \leq Mh_l \quad \forall l \in L, \quad (14)$$

$$\sum_{ij \in A_r : i < j} x_{ij}^l \geq h_l \quad \forall l \in L, \quad (15)$$

$$h_l + \sum_{ij \in A_r : i < j} x_{ij}^l = \sum_{i \in N_r} y_i^l \quad \forall l \in L. \quad (16)$$

Constraints (14) ensure that an arc ij is assigned to line l if and only if line l is located; here, M is an upper bound to the number of arcs assigned to lines. Constraints (15) say that a line l is located if and only if at least an arc ij is assigned to it. Constraints (16) say that lines cannot have assigned more stations than the number of assigned arcs minus one.

3.3.3. Demand constraints.

$$\sum_{k \in N(i)} f_{ki}^w - \sum_{j \in N(i)} f_{ij}^w = \begin{cases} -1 & \text{if } i = o(w) \quad \forall w \in W, \\ 1 & \text{if } i = d(w) \quad \forall w \in W, \forall i \in N, \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

$$\sum_{ij \in A_r \cup A_d \cup A_o} d_{ij} f_{ij}^w \leq \mu_w u_{\text{cur}}^w (1 - f_{\text{cur}}^w) \quad \forall w \in W. \quad (18)$$

Constraints (17) are multi-commodity flow conservation constraints at each node and for each passenger group. They state that every passenger group must enter (in the origin centroid), leave (in the destination centroid) or continue (in the rest of nodes) in the network. Constraints (18) and the objective function determine for each passenger group w the shortest path between the new network and the current network. On the one hand, if $f_{\text{cur}}^w =$

0, then there is a path for passenger group w in the new RTN shorter (i.e. cheaper) than or equal to $\mu_w u_{\text{cur}}^w$. On the other hand, if $f_{\text{cur}}^w = 1$, then $f_{ij}^w = 0$ for all arcs $(i, j) \in A_r \cup A_d \cup A_o$.

3.3.4. Location–allocation constraints.

$$f_{ij}^w + f_{ji}^w \leq \chi_{ij} \quad \forall ij \in A_r : i < j, \forall w \in W, \quad (19)$$

$$f_{o(w)j}^w \leq \psi_j \quad \forall j \in N_r, \forall w \in W, \quad (20)$$

$$f_{id(w)}^w \leq \psi_i \quad \forall i \in N_r, \forall w \in W, \quad (21)$$

Constraints (19)–(21) say that each passenger group w can make use of arcs ij and nodes i and j if and only if those infrastructures are located in the new RTN.

4. RRTND model

The previous RTND model does not consider the capacity of the system: we assume that the public transportation costs are independent of arcs' passenger flow, that is, the available capacity is enough. In this context, we assume the all-or-nothing assignment for passenger demand. Therefore, all the passengers in a passenger group will commute using the same path. And consequently, some of the arcs in the network will feature great traffic levels while other will feature low traffic levels. This is not desirable from the robustness point of view: a disruption, such as an arc closure, may involve a large number of passengers.

Laporte et al. (2011) propose a model for the design of a robust RTN. They state that a network is said to be robust when the effects of disruptions on total trip coverage is minimised. The proposed model is constrained by three different kinds of flow conditions: demand-arc flow constraints, arc flow constraints and arc demand constraints. These constraints yield a network that provides several alternative routes for given origin–destination pairs, therefore increasing robustness. For example, constraints (22) are demand-arc flow constraints.

$$f_{ij}^w \leq \frac{1}{q_{ij}^w} \quad \forall ij \in A_{\text{dof}} \subseteq A, w \in W_{\text{dof}} \subseteq W, \quad (22)$$

where A_{dof} and W_{dof} are subsets of arcs and passenger groups, respectively. q_{ij}^w is the robustness parameter, that is, the number of paths in which the passenger group must be split on. Consequently, the variables modelling the passenger flow on the arcs must be relaxed:

$$f_{ij}^w \in [0, 1] \quad \forall ij \in A_{\text{dof}}, w \in W_{\text{dof}}. \quad (23)$$

Constraints (22) account for demand routing patterns and ensure that for a given arc $ij \in A_{\text{dof}}$ the percentage of a passenger group $w \in W_{\text{dof}}$ using it must be lower than $1/q_{ij}^w$ (i.e. passengers from the same passenger group must be routed through different paths). This multi-path routing approach splits up the demand over several arc-disjoint paths. These paths are the q arc-disjoint shortest paths for each passenger group (see Laporte et al. 2011 for more details). The difference between the objective function values in the RTND model and the RRTND model is the price of robustness. The RRTND model is defined by the objective function in Equation (1) and constraints (5)–(23).

The robust model presented here does not need to account for disrupted scenarios explicitly. This information is stored in A_{dof} , W_{dof} and q_{ij}^w . A_{dof} is the subset of arcs where

we expect to have disruptions, W_{daf} is the subset of passenger groups we want to ensure they have multiple paths available to reach destination and q_{ij}^w is the number of paths we want to have for each passenger group in W_{daf} and arc in A_{daf} . That is, we can enforce paths which are totally different from origin to destination, or only different at certain critical arcs.

This modelling approach does not only provide alternative paths in the event of disruptions, but also avoids the consolidation of traffic in a few arcs, and therefore congestion.

5. Recoverable RRTND model

Traditional robust approaches do not consider that solutions may be recovered after data perturbation, that is, after disruption occurrence. Therefore, the required immunity may be less strict if the recovery of the system is considered.

Recoverable robustness explicitly takes into account available recovery actions once the disruption has occurred: a recoverable robust solution does not need to be feasible for a whole set of admissible disturbances, but it must be able to be recovered from a set of disrupted scenarios.

The rest of this section is divided into different subsections. The first one is devoted to disruption modelling approach. The second one to the recovery. And the last one, to the mathematical model.

5.1. Disruptions

We assume that disruptions are characterised by future arc malfunction/closure due to causes such as vehicle breakdowns, infrastructure malfunction, collisions and outrages. We acknowledge that there are disruptions of different nature (e.g. earthquakes vs. system malfunction); however, we focus on disruptions which only affect the new RTN. We also assume that only a subset of the arcs will fail during the operation of the new RTN: they will likely be the arcs with the greatest traffic previsions.

We will consider a set S of admissible disruptions. Each scenario $s \in S$ will represent a different disruption. Disruptions may be modelled by assigning extremely high routing costs (i.e. penalties) to the arcs where the disruption has occurred. Consequently, these arcs will be never used by passengers: they will find cheaper paths in the new RTN or in the current network. In the mathematical model in Section 3, d_{ij} and u_{cur}^w are the distance of each arc ij and the distance for passenger group w in the current network, respectively. Now, d_{ij} is different for each scenario $s \in S$. Whenever a disruption hits the new network, d_{ij}^s is increased with respect their nominal values, respectively. Passengers will decide their path according to the costs at each scenario $s \in S$.

5.2. Recovery

Recall we need a recovery strategy in order to recover the system from each different disruption. As we are dealing with a long term network design problem, we must be aware that it is not possible to construct new arcs when a disruption hits the network in order to recover from it. Therefore, the recovery strategy is to re-schedule passengers' flows by modelling their path selection under the new network topology. This new network topology will be the original network but with those arcs affected by the disruption closed (i.e.

by assigning them an extremely high distance or routing cost). We assume that a line with a disrupted arc continues operating in the other arcs. This is consistent with the fact that we are deciding here the core network and secondary stations are to be decided later (Laporte et al. 2007). Therefore, we assume that there will be recovery tracks or switches to change tracks in all the stations considered here in such a way that a line with a disrupted arc can still be operated in the other arcs.

We need to define the following new passenger flow variables:

- $f_{ij}^{w,s} \in \{0, 1\}$ equals 1 if passenger group w in scenario s uses arc ij in the new network; 0, otherwise.
- $f_{cur}^{w,s} \in \{0, 1\}$ equals 1 if passenger group w in scenario s uses the current network; 0, otherwise.

5.3. Mathematical model

The Recoverable Robust RTND Model (RR-RTNDM) is a two-stage programming model where decisions may be split into two different sets: the first set of decisions is made before disruption's information realisation (i.e. data revelation regarding the disruption), and the second one is made once uncertainty is solved (i.e. once we know all the details of the disruptions):

- The first stage considers the non-disrupted scenario: the network design variables $x_{ij}^l, \chi_{ij}, y_i^l, \psi_i, h_l$ and the passenger variables f_{ij}^w, f_{cur}^w ; the constraints are given by Equations (5)–(21).
- The second stage considers passenger flows at each disruption, which are modelled after disruption realisation; passenger variables will be $f_{ij}^{w,s}, f_{cur}^{w,s}$ and constraints (17)–(21) will be reproduced for each different disruption.

Each disruption's occurrence probability is not known. Therefore, the two-stage RR-RTNDM objective is to cope with uncertainty by minimising the maximum negative effects produced by a disruption in the set of disruptions. The RR-RTNDM aims at minimising the sum of the nominal costs (i.e. costs for an undisturbed scenario) and the maximum recovery costs over the set of disrupted scenarios. Here, recovery costs are modelled as the number of passengers using the current network at each disruption. Consequently, this approach relieves us from using probabilities for different disruption occurrences.

We define Λ_s , which is a positive variable, to determine the number of passengers using the current network at each disruption:

$$\Lambda_s = \sum_{w \in W} g_w f_{cur}^{w,s} \quad \forall s \in S. \quad (24)$$

Therefore, the maximum number of passengers using the current network within the set of disruptions will be

$$\Lambda = \max_{s \in S} \{\Lambda_s\} = \max_{s \in S} \left\{ \sum_{w \in W} g_w f_{cur}^{w,s} \right\}, \quad (25)$$

where Λ is again a positive variable.

The RR-RTNDM may be briefly described as follows:

$$\min \quad z = \alpha z_{\text{cur}} + \beta z_{\text{loc}} + \gamma z_{\text{route}} + \alpha \Lambda \quad (26)$$

subject to :

$$A\mathbf{x} \geq b, \quad (27)$$

$$E\mathbf{y} - U\mathbf{x} \geq d, \quad (28)$$

$$E_s \mathbf{y}_s - U\mathbf{x} \geq d \quad \forall s \in S, \quad (29)$$

$$\Lambda = f(\mathbf{y}_1, \dots, \mathbf{y}_{|S|}), \quad (30)$$

$$\mathbf{x} \in \{0, 1\}^{ijl+ij+il+i+l}, \quad (31)$$

$$\mathbf{y} \in \{0, 1\}^{ijw+w}, \quad (32)$$

$$\mathbf{y}_s \in \{0, 1\}^{ijw+w} \quad \forall s \in S, \quad (33)$$

$$\Lambda \geq 0 \quad (34)$$

where we use the same weight for Λ and z_{cur} in the objective function (i.e. both terms represent passengers attracted to the current network), $\mathbf{x} = \{x_{ij}^l, \chi_{ij}, y_i^l, \psi_i, h_l\}^T$ is the vector of design variables, $\mathbf{y} = \{f_{ij}^w, f_{\text{cur}}^w\}^T$ is the vector of passenger flow variable for the undisrupted situation, $\mathbf{y}_s = \{f_{ij}^{w,s}, f_{\text{cur}}^{w,s}\}^T$ is the vector of passenger flow variables for each of the disrupted scenarios, constraints (27) represent constraints (5)–(16), constraints (28) represent constraints (17)–(21), constraints (29) are passenger flow constraints for each scenario, and Equation (30) calculates the maximum number of passengers using the current network within the set of disrupted scenarios (Equation (25)).

The RR-RTNDM given by Equations (26)–(34) is a nonlinear mixed-integer programming model. The following formulation introduces in detail block of constraints (29) and linearises Equation (30):

$$\min \quad z = \alpha z_{\text{cur}} + \beta z_{\text{loc}} + \gamma z_{\text{route}} + \alpha \Lambda \quad (35)$$

$$\text{subject to : } A\mathbf{x} \geq b, \quad (36)$$

$$E\mathbf{y} - U\mathbf{x} \geq d, \quad (37)$$

$$\sum_{k \in N(i)} f_{ki}^{w,s} - \sum_{j \in N(i)} f_{ij}^{w,s} = \begin{cases} -1 & \text{if } i = o : w = (o, d) \in W \\ 1 & \text{if } i = d : w = (o, d) \in W \\ 0 & \text{otherwise} \end{cases}$$

$$\forall i \in N, \forall w \in W, \forall s \in S, \quad (38)$$

$$\sum_{ij \in A_r \cup A_d \cup A_o} d_{ij}^s f_{ij}^{w,s} \leq \mu_w u_{\text{cur}}^w (1 - f_{\text{cur}}^{w,s}) \quad \forall w \in W, \forall s \in S, \quad (39)$$

$$f_{ij}^{w,s} + f_{ji}^{w,s} \leq \chi_{ij} \quad \forall ij \in A_r, \forall w \in W, \forall s \in S, \quad (40)$$

$$f_{o(w)j}^{w,s} \leq \psi_j \quad \forall j \in N_r, \forall w \in W, \forall s \in S, \quad (41)$$

$$f_{id(w)}^{w,s} \leq \psi_i \quad \forall i \in N_r, \forall w \in W, \forall s \in S, \quad (42)$$

$$\Lambda \geq \sum_{w \in W} g_w f_{\text{cur}}^{w,s} \quad \forall s \in S, \quad (43)$$

$$\mathbf{x} \in \{0, 1\}^{ijl+ij+il+i+l}, \quad (44)$$

$$\mathbf{y} \in \{0, 1\}^{ijw+w}, \quad (45)$$

$$\mathbf{y}_s \in \{0, 1\}^{ijw+w} \quad \forall s \in S, \quad (46)$$

$$\Lambda \geq 0 \quad (47)$$

Constraints (38)–(42) are the block of constraints (29): they model passenger flows at each disrupted scenario. Constraints (43) calculate the maximum number of passengers using the current network within the set of disrupted scenarios.

6. Computational experiments

Our experiments are based on two different cases: network R1 and Seville's rail RTN (Escudero and Muñoz 2009; Laporte et al. 2007; Marín 2007). Network R1 has 9 potential stations, 30 potential arcs, 72 passenger groups and a total demand of 1044 passengers. Each potential station i in network R1 has an associated construction cost c_i and each potential arc ij a pair (c_{ij}, d_{ij}) of weights: the construction cost c_{ij} and the distance d_{ij} , which is used to define the generalised trip cost. Figure 2 depicts the potential network R1.

In network R1, the origin–destination demand g_w for each passenger group $w \in W$ is defined by the following matrix:

$$G = \begin{pmatrix} & 9 & 26 & 19 & 13 & 12 & 13 & 8 & 11 \\ 11 & - & 14 & 26 & 7 & 18 & 3 & 6 & 12 \\ 30 & 19 & - & 30 & 24 & 8 & 15 & 12 & 5 \\ 21 & 9 & 11 & - & 22 & 16 & 25 & 21 & 23 \\ 14 & 14 & 8 & 9 & - & 20 & 16 & 22 & 21 \\ 26 & 1 & 22 & 24 & 13 & - & 16 & 14 & 12 \\ 8 & 6 & 9 & 23 & 6 & 13 & - & 11 & 11 \\ 9 & 2 & 14 & 20 & 18 & 16 & 11 & - & 4 \\ 8 & 7 & 11 & 22 & 27 & 17 & 8 & 12 & - \end{pmatrix}.$$

The generalised cost u_{cur}^w for each passenger group $w \in W$ in the current network is defined by the matrix:

$$U^{\text{CUR}} = \begin{pmatrix} - & 1.6 & 0.8 & 2 & 1.6 & 2.5 & 3 & 2.5 & 0.8 \\ 2 & - & 0.9 & 1.2 & 1.5 & 2.5 & 2.7 & 2.4 & 1.8 \\ 1.5 & 1.4 & - & 1.3 & 0.9 & 2 & 1.6 & 2.3 & 0.9 \\ 1.9 & 2 & 1.9 & - & 1.8 & 2 & 1.9 & 1.2 & 2 \\ 3 & 1.5 & 2 & 2 & - & 1.5 & 1.1 & 1.8 & 1.7 \\ 2.1 & 2.7 & 2.2 & 1 & 1.5 & - & 0.9 & 0.9 & 2.9 \\ 2.8 & 2.3 & 1.5 & 1.8 & 0.9 & 0.8 & - & 1.3 & 2.1 \\ 2.8 & 2.2 & 2 & 1.1 & 1.5 & 0.8 & 1.9 & - & 0.3 \\ 1 & 1.5 & 1.1 & 2.7 & 1.9 & 1.8 & 2.4 & 3 & - \end{pmatrix}.$$

The second network is based on Seville's city in Spain. It has 24 potential stations, 264 potential arcs, 552 passenger groups and a total demand of 292,000 passengers. This potential network is graphed in Figure 3. Because this is a realistic network we need to identify

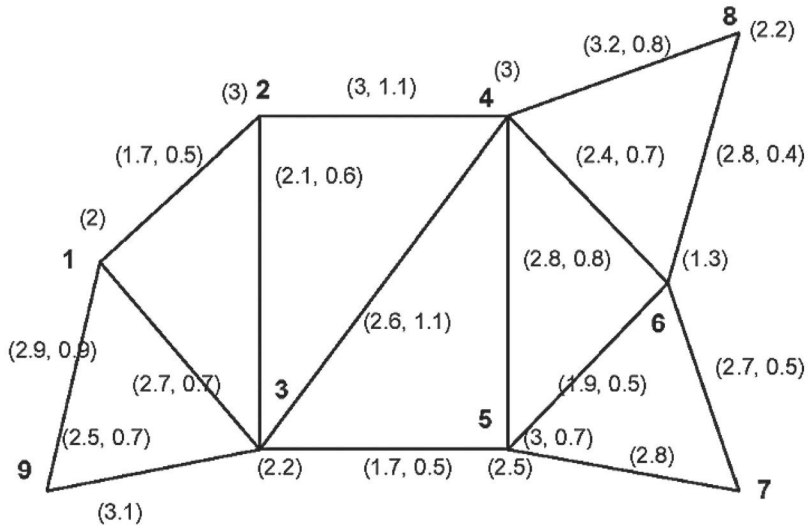


Figure 2. Network R1.

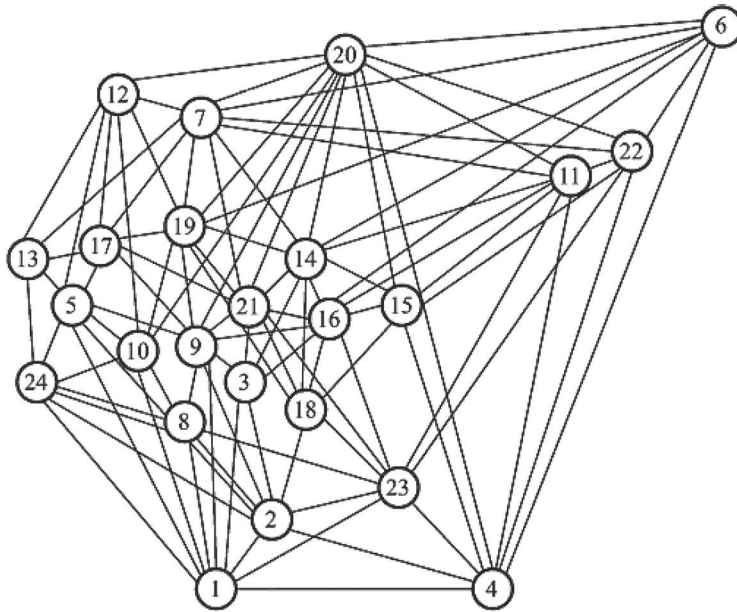


Figure 3. Seville's network.

potential stations and arcs. For this purpose, we follow the process in Laporte et al. (2007): the area under consideration is partitioned into zones, according to the number of trips that each zone will produce or attract and different potential stations are identified; then, all the possible arcs are identified. See Laporte et al. (2007) for more details on this.

We present two different studies in the following subsections. In Section 6.1, we evaluate our new recoverable robust approach. We show how it produces robust solutions which

can be recovered at a limited cost. In Section 6.2, we perform a multiobjective optimisation study.

For all the case studies in this paper, we need to study disrupted scenarios. As explained in Section 5 we assume that disruptions are characterised by arc closures. Therefore, we consider a set of disruptions, where each disruption is produced by a different arc closure. These arcs are the arcs with the greatest traffic previsions. This set is ordered according to these traffic previsions, that is, the arc with the greatest traffic prevision is the first element of the set, the arc with second greatest traffic prevision is the second element and so on.

We used for our tests a personal computer with an Intel Core i7 at 2.8 GHz and 8 GB of RAM, running under Windows 7 64-Bit, and we implemented the models in GAMS/Cplex 12.1.

6.1. Evaluation of the recoverable robust approach

We present the computational results regarding network R1 and Seville's network in a sequential fashion in this section. For the computational experiments, we use the values for the weights in the objective function given by Marín and Jaramillo (2008), where α is a number close to 1, because the new network trip covering is the main component of the objective function. The other terms (β , γ) are included to adequately simulate the routing user behaviour and the location of any facility, which is not free of cost ($\beta = (1 - \alpha)/2$, $\gamma = (1 - \alpha)/2$).

One of the key parameters in the RR-RTND mathematical model presented before is the number of scenarios to be included. Obviously, this number will influence the solution. We need to make sure that the number of scenarios we include in the mathematical model is representative enough of the real situation.

Figures 4 and 5 show the evolution of the attended demand by the new network for the worst-case and the RR-RTND objective function value depending on the number of scenarios included for the networks in Figures 2 and 3, respectively. We define the worst-case as the scenario which provides the greatest value for Δ_s . In order to obtain these evolutions, we solve the RR-RTND model as many times as scenarios (i.e. number of disruptions) we have. The number of scenarios ranges from 0 to a positive integer number. 0 corresponds to the solution where no disruptions are considered, that is, no disruption hits the network and the network operates normally. 1 corresponds to the solution where the disruption related to the arc with the greatest traffic prevision is considered. 2 corresponds to the solution where the disruptions related to the two arcs with the greatest traffic previsions are considered, and so on. Therefore, for each case we obtain different values for the attended demand and the objective function. When the number of scenarios included in the mathematical model is low, the RR-RTND model produces solutions which are able to attend a big number of passenger demand, that is, the number of passengers that commutes through the current network is low. However, the number of passengers that uses the new network decreases when the number of scenarios increases and it reaches a constant value regardless of the number of scenarios.

When the curves for the attended demand and the objective function reach a constant value, it means that the addition of more scenarios has no influence in the solution; typically the rest of the disruptions feature arcs which have lower levels of traffic (i.e. if the ordering of the set of disruptions is correct). This means that, given a critical number of

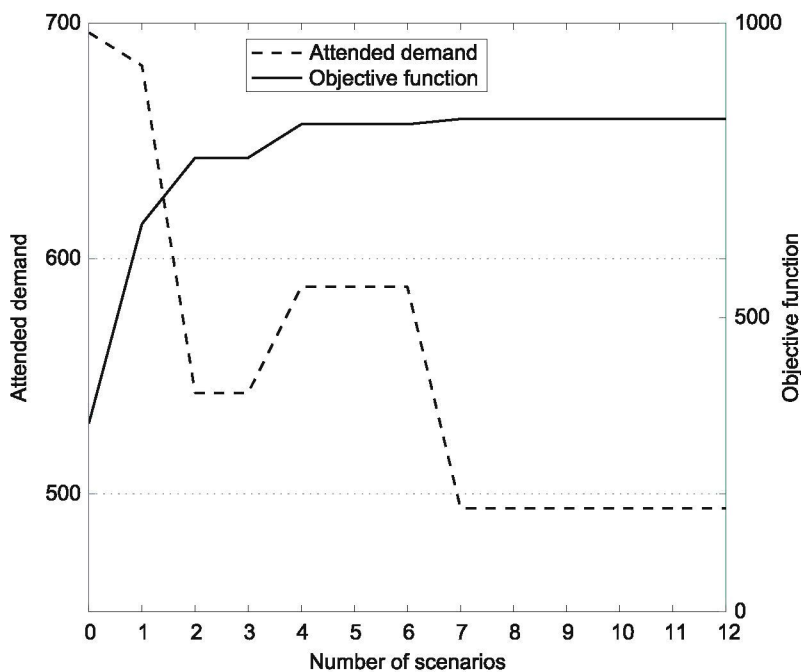


Figure 4. Evolution of the attended demand for the worst-case and the RR-RTND objective function value depending on the number of scenarios for network R1.

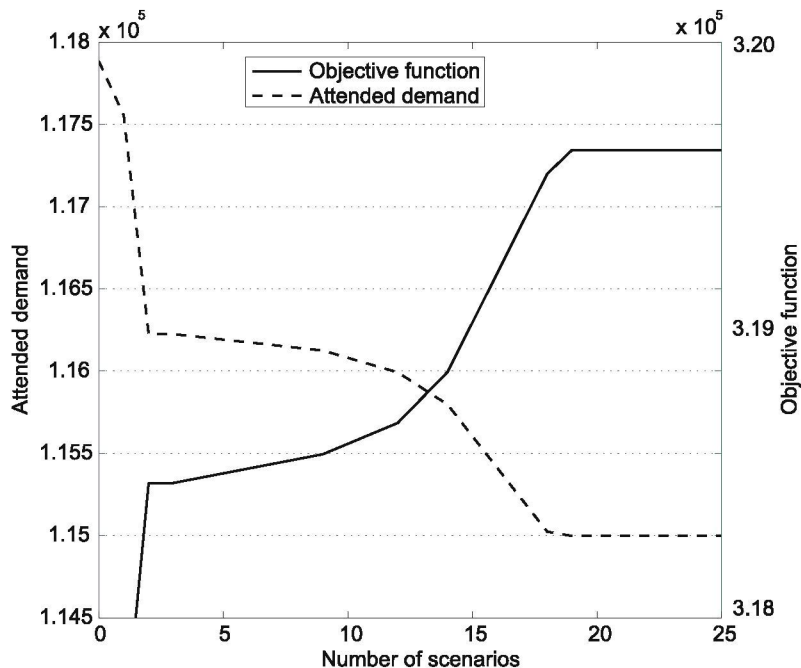


Figure 5. Evolution of the attended demand for the worst-case and the RR-RTND objective function value depending on the number of scenarios for Seville's network.

Table 1. RR-RTND model size for network R1 and 7 scenarios.

Item	RR-RTND
Number of discrete variables	29,626
Number of continuous variables	4
Number of constraints	27,411
Number of non-zero elements	110,225

Table 2. RR-RTND model size for Seville's network and 19 scenarios.

Item	RR-RTND
Number of discrete variables	3,127,319
Number of continuous variables	4
Number of constraints	2,122,916
Number of non-zero elements	12,858,393

scenarios, the RR-RTNDM produces a network design which can be recovered from them at a maximum cost which is in fact minimum, that is, the maximum value of Δ which is the number of passengers that use the current network is minimised across all the scenarios. For these case studies, the critical number of scenarios to be included in the model in order to have a enough representative set of disruptions are 7 and 19, respectively. We choose these numbers of scenarios because adding more scenarios requires more computational efforts without added value to the solutions.

Tables 1 and 2 show the mathematical model size for the network R1 and 7 scenarios and Seville's network and 19 scenarios, respectively. The RR-RTNDM numbers of discrete and continuous variables, constraints and non-zero elements are given.

In order to compare the three approaches presented in this paper (i.e. the RTNDM, the RRTNDM and the RR-RTNDM), Figures 6 and 7 show the deviation value for each of the scenarios and approach for network R1 and Seville's network, respectively. We define the deviation for each scenario as $\Delta_s = \sum_{w \in W} g_w(f_{cur}^{w,s} - f_{cur}^w)$. Recall that the aim of the RR-RTNDM is to minimise the maximum value of Δ (i.e. $\min \Delta = \min_{s \in S} \max_{w \in W} g_w f_{cur}^{w,s}$). There is one vertical bar for each of the scenarios and approaches representing Δ_s (when there is no vertical bar for an approach, it means that $\Delta_s = 0$): black bars for the RR-RTNDM solutions, dark grey bars for the RRTNDM solutions and grey bars for the RTNDM solutions. The bars are always displayed in the same order. Each of them represents the number of passengers that cannot travel within the new network due to the disruption as compared to the case where no disruption hits the network, that is, Δ_s .

For a deeper insight in the results, Tables 3 and 4 provide some metrics for network R1 and Seville's network, respectively. For each of the solutions (i.e. RTND, RRTND and RR-RTND solutions) Tables 3 and 4 show the following columns: the second one shows the location cost (z_{loc}), the third one the demand covered by the current network (z_{cur}), the fourth column the routing cost in the new network (RCN), the fifth one the maximum number of passengers using the current network (Δ), the sixth one the mean value of all the Δ_s values (μ_Δ), the seventh one the standard deviation of all the Δ_s values (σ_Δ) and the last column the solution time in seconds. Note that RCN is the routing cost in the new network, while z_{route} is the total routing cost, that is, the sum of the routing cost in the new and in the current networks.

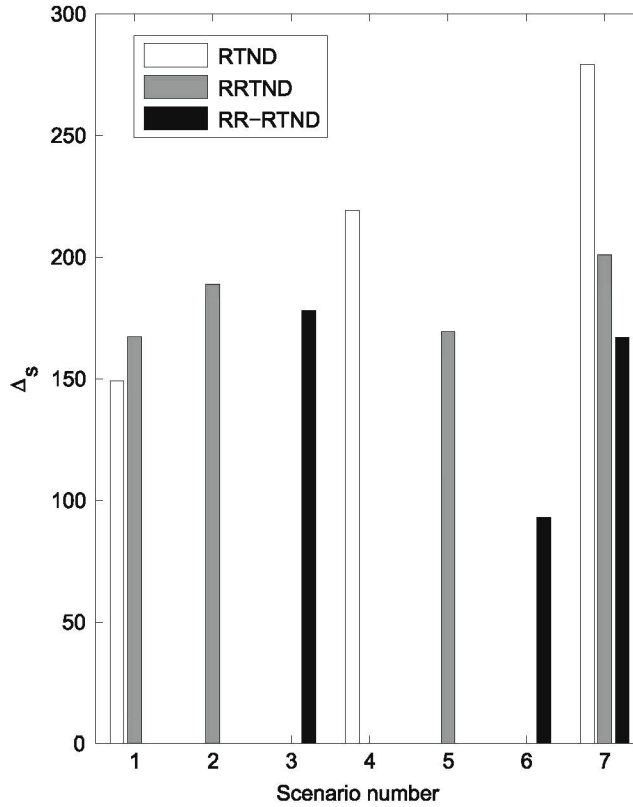


Figure 6. Deviation value for each of the scenarios and approach for network R1 (Δ_s).

The location cost is similar for all the solutions. However, the attended demand varies. We can see how the RRTND solution is the worst one in terms of attended demand: many trips are covered by the current network. This means that the price of robustness remains on the number of attended passengers. RTND and RR-RTND show similar behaviour in terms of attended demand. The total routing cost is proportional to the total distance travelled for each origin–destination pair. Therefore, the RCN increases as the number of passengers attended by the current network decreases. The maximum number of passengers using the current network (Δ) is a key indicator for the quality of the solution. It tells us how the solution behaves under the worst-case scenario (among all the considered scenarios), the greater value, the poorer performance. The solution that performs worst is always the RTND solution. But we cannot only focus on the worst-case. It is also important to know how the solution behaves across the different scenarios. Additional to the visual display in Figures 6 and 7, the mean and standard deviation values of all the Δ_s values give information about the solution behaviour. For network R1, the lowest values are again obtained for the RR-RTND approach; therefore, the RR-RTND solution will perform better than the other two solutions across the different scenarios. However, for Seville's network the lowest values arise for the RRTND solution; but we must note that this solution is significantly worse in terms of attended demand: more than 50,000 passengers are lost to the current network.

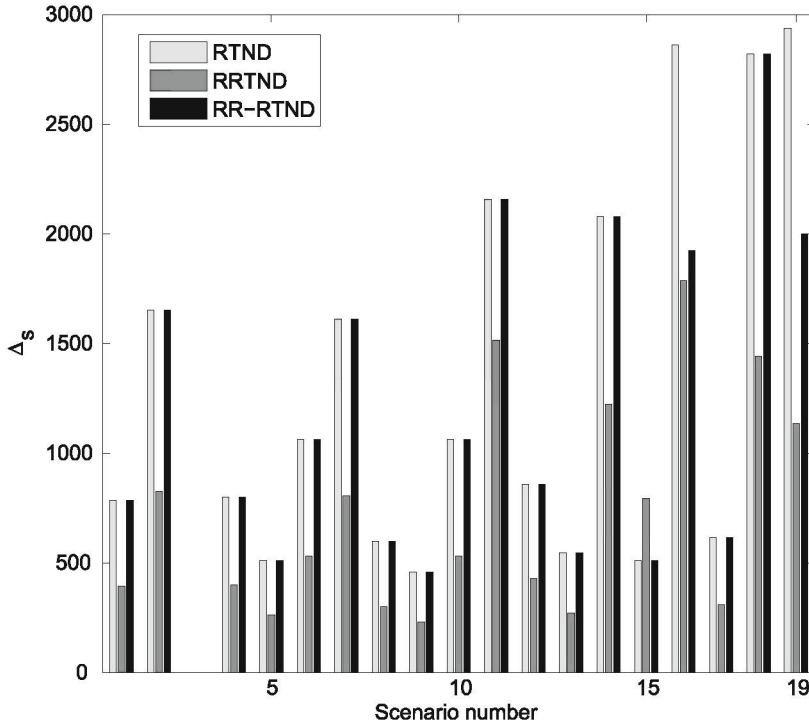


Figure 7. Deviation value for each of the scenarios and approach for Seville's network (Δ_s).

Table 3. RTND, RRTND and RR-RTND solutions for network R1.

Solution	z_{loc}	z_{cur}	RCN	Δ	μ_{Δ}	σ_{Δ}	ST
RTND	49.60	348	60.80	283	92.42	121.24	0.38
RRTND	48.60	619.61	47.08	200.89	103.79	97.75	1.94
RR-RTND ^a	49.50	372	45.50	178	62.57	82.47	3.59

^a Solution for 7 scenarios.

Table 4. RTND, RRTND and RR-RTND solutions for Seville's network.

Solution	z_{loc}	z_{cur}	RCN	Δ	μ_{Δ}	σ_{Δ}	ST
RTND	9997.69	175,022	7532.60	2936	1259.40	911.50	65.02
RRTND	9999.13	230,714.68	3996.04	1786.5	693.78	505.84	113.11
RR-RTND ^a	9997.60	175,088	7517.30	2821	1160.70	759.49	8390.67

^a Solution for 19 scenarios.

6.2. Multiobjective optimisation

This subsection presents two different multiobjective optimisation case studies: the first one regarding network R1, and the second one regarding Seville's network.

The model presented before aims at producing an efficient network design while enabling efficient recovery solutions against disruptions. These two aims are usually contradictory objectives: efficient network designs usually minimise resources allocation-location while recovery plans which minimise negative disruptive effects usually need and spend more resources. We are interested in the trade-off between efficient and recoverable robust

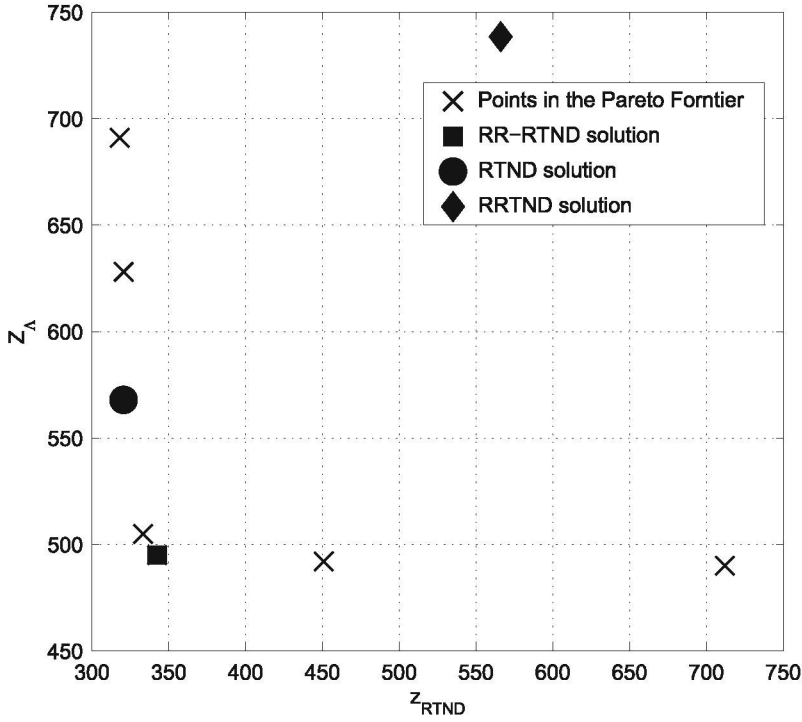


Figure 8. Pareto frontier for network R1.

network design solutions. Therefore, we propose a new objective function in order to make a bi-objective study as follows:

$$z = \eta(\alpha z_{cur} + \beta z_{loc} + \gamma z_{route}) + (1 - \eta)(\alpha \Lambda) = \eta z_{RTND} + (1 - \eta)(z_{\Delta}). \quad (48)$$

The bi-objective function is subject to constraints (36)–(47). Here $\eta \in [0, 1]$. Different values of η may yield different Pareto solutions for the new bi-objective optimisation problem. z_{RTND} and z_{Δ} represent the two objective numerical values. The rest of the parameters take the same value as in the previous subsection.

Figures 8 and 9 show the Pareto frontiers for the case studies drawn from network R1 with 7 scenarios and Seville's network with 19 scenarios, respectively. The x-axes represent z_{RTND} and the y-axes represent z_{Δ} . Because we are dealing with a mixed-integer linear programming model the Pareto frontier is not continuous

For a comparison, Figures 8 and 9 also plot the RTND (the circle), RRTND (the rhombus) and RR-RTND (the square) solutions.

The RTND and RR-RTND solutions turn out to be Pareto optimal; however, the RRTND solution is far way from the Pareto frontier (for Seville's network it is out of the scale: it is the point given by (209572,209251)). The RTND solution features the lowest possible z_{RTND} value, that is, the lowest value for the combination of passengers lost to the current network and location and routing costs. However, the RR-RTND does account for the z_{Δ} term and it produces a solution which is a reasonable trade-off between the two objectives.

The modified objective (48) is computationally more challenging than the original objective (35). We experienced solution times for network R1 and Seville's network up to 85.28

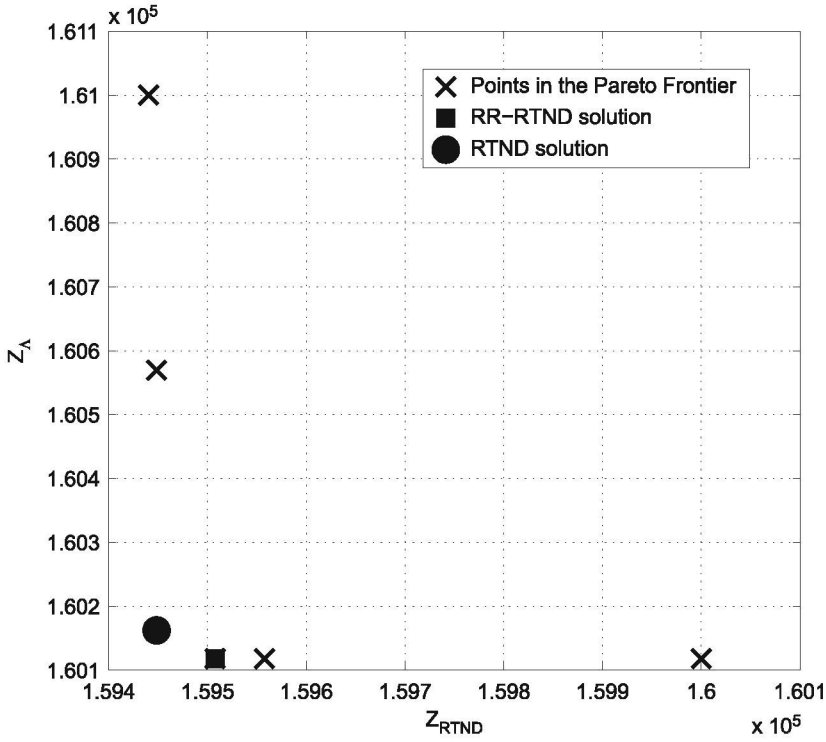


Figure 9. Pareto frontier for Seville's network.

and 24152 seconds, respectively, mostly in cases where the demand deviation cost value (z_Δ) is penalised more heavily. In contrast, the RR-RTND solution needed 3.59 and 8390.67 s, respectively.

6.3. Summary of the computational results

The proposed algorithmic framework allows us to find recoverable robust solutions to the RTND problem. The recoverable robust solutions are better than those obtained with previously developed approaches. We have presented two case studies, which are drawn from network R1 and Seville's network. The obtained results are virtually the same and may be summarised as the following two bullet points:

- we have demonstrated how the robustness concept presented in the RRTND approach fails to produce acceptable solutions. Indeed, it produces robust solutions, that is, Δ , μ_Δ and σ_Δ are generally ameliorated, but their price in terms of attended demand is huge: the RRTND solutions for network R1 and Seville's network lose a 62.47% and a 83.17% of the demand to the current network, respectively;
- and we have demonstrated how the recoverable robust concept presented in the RR-RTND approach produces a great trade-off between robustness and efficiency: while it produces solutions acceptably close to the ones produced by the RTND in terms of cost and attended demand, Δ , μ_Δ and σ_Δ are ameliorated: a 37.1%, 32.29% and

31.97% for network R1, respectively, and a 3.91%, 7.83% and 16.67% for Seville's network, respectively.

7. Conclusions

We have presented a recoverable robust rapid transit railway network design problem in this paper. The presented approach aims at finding a network design emphasising the trade-off between efficiency and robustness accounting for the possibility of recovering from potential disruptions. That is, the network design does not need to be feasible for all the disruptions, but it must be able to be recovered to get a feasible solution.

We have developed a new formulation which addresses the minimisation of the effort that needs to be made so as to recover from disruptions while maximising passenger trip coverage. This formulation is based on previous research related to non-robust and RRTND. The major novelty of the proposed methodology is in the mini-max approach of disruption costs, which aims at limiting the negative disruptive impacts in terms of passengers that have to use the current network, and are lost by the new one in disrupted scenarios. We also present an approach in order to select the number of disrupted scenarios to be included in the modelling approach. Moreover, we have compared the solutions obtained by the three approaches, namely the non-robust, the robust and the recoverable RRTND approaches, in order to show that non-RRTND is too optimistic and that RRTND is extremely conservative as compared to the recoverable robust approach presented in this paper.

Computational experiments drawn from fictitious and realistic networks, which are conducted within reasonable computational times, show how the robustness concept presented in traditional robust approaches fails to produce acceptable solutions and the recoverable robust approach produces a great trade-off between robustness and efficiency. The results show that the proposed methodology performs better than the traditional one in terms of recovery costs. Compared to the robust approach, the recoverable robust one has a larger recovery cost. But the number of passengers attracted to the new network is significantly larger for the recoverable robust model, making it more suitable.

Acknowledgements

Thanks are due to the four anonymous reviewers for their useful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by Project Grants TRA2011-27791-C03-01 and TRA2014- 52530-C3-1-P by the 'Ministerio de Economía y Competitividad', Spain.

References

- Ben-Tal, A., and A. Nemirovski. 1999. "Robust Solutions of Uncertain Linear Programs." *Operations Research Letters* 25: 1–13.
- Ben-Tal, A., and A. Nemirovski. 2002. "Robust Optimization-Methodology and Applications." *Mathematical Programming, Series B* 92: 453–480.

- Bertsimas, D., and M. Sim. 2003. "Robust Discrete Optimization and Network Flows." *Mathematical Programming* 98: 49–71.
- Bertsimas, D., and M. Sim. 2004. "The Price of Robustness." *Operations Research* 52: 35–52.
- Bruno, G., M. Gendreau, and G. Laporte. 2002. "A Heuristic for the Location of a Rapid Transit Line." *Computers and Operations Research* 29: 1–12.
- Cacchiani, V., A. Caprara, L. Galli, L. Kroon, G. Maróti, and P. Toth. 2012. "Railway Rolling Stock Planning: Robustness Against Large Disruptions." *Transportation Science* 46 (2): 217–232.
- Cadarso, L., and Á. Marín. 2012a. "Integration of Timetable Planning and Rolling Stock in Rapid Transit Networks." *Annals of Operations Research* 199 (1): 113–135.
- Cadarso, L., and Á. Marín. 2012b. "Robustness in Rapid Transit Network Design." *Procedia Social and Behavioral Sciences* 54: 1288–1297.
- Cadarso, L., Á. Marín, and G. Maróti. 2013. "Recovery of Disruptions in Rapid Transit Networks." *Transportation Research Part E: Logistics and Transportation Review* 53: 15–33.
- Cadarso, L., and Á. Marín. 2015. "Improved Rapid Transit Network Design Model: Considering Transfer Effects." *Annals of Operations Research*. doi:10.1007/s10479-015-1999-x.
- Cadarso, L., G. Maróti, and Á. Marín. 2015. "Smooth and Controlled Recovery Planning of Disruptions in Rapid Transit Networks." *IEEE Transactions on Intelligent Transportation Systems* 16 (4): 2192–2202.
- Caprara, A., L. Galli, S. Stiller, and P. Toth. 2008. *Recoverable-robust Plataforming by Network Buffering*. Technical Report ARRIVAL-TR.0157. ARRIVAL project <http://www.cti.gr/arrival>.
- Cicerone, S., G. D'Angelo, G. Di Stefano, D. Frigioni, A. Navarra, M. Schachtebeck, and A. Schöbel. 2009. "Recoverable Robustness in Shunting and Timetabling." In *Robust and On-Line Large Scale Optimization*, edited by R. Ahuja, R. Möhring, and C. Zaroliagis, 28–60, Lectures Notes in Computer Science, Vol. 5868. Berlin: Springer-Verlag.
- Codina, E., A. Marín, and L. Cadarso. 2014. "Robust Infrastructure Design in Rapid Transit Rail Systems." *Transportation Research Procedia* 3: 660–669.
- Curtin, K. M., and S. Biba. 2011. "The Transit Route Arc-Node Service Maximization Problem." *European Journal of Operational Research* 208: 46–56.
- Eggenberg, N. 2009. "Combining Robustness and Recovery for Airline Schedules." PhD thesis, Ecole Polytechnique Federale De Lausanne.
- Escudero, L. F., and S. Muñoz. 2009. "An Approach for Solving a Modification of the Extended Rapid Transit Network Design Problem." *Top* 17 (2): 320–334.
- Fischetti, M., D. Salvagin, and A. Zanette. 2011. "Fast Approaches to Improve the Robustness of Railway Timetable." *Transportation Science* 43 (3): 321–335.
- Froyland, G., S. J. Mahen, and Ch.-L. Wu. 2014. "The Recoverable Robust Tail Assignment Problem." *Transportation Science* 49 (3): 351–372.
- García, R., A. Marín, J. A. Mesa, F. Perea, and D. Verasategui. 2007. "A New Concept of Robustness." Proceedings of ATMOS, Seville, Spain, November 15–16, 1–14.
- Grönkvist, M. 2005. "The Tail Assignment Problem." PhD thesis, Chalmers University of Technology and Gutenberg University.
- Gutiérrez-Jarpa, G., C. Obreque, G. Laporte, and V. Marianov. 2013. "Rapid Transit Network Design for Optimal Cost and Origin-Destination Demand Capture." *Computers & Operations Research* 40: 3000–3009.
- Laporte, G., Á. Marín, J. A. Mesa, and F. A. Ortega. 2007. "An Integrated Methodology for the Rapid Transit Network Design Problem. Lecture Notes in Computer Science. *Algorithmic Methods for Railway Optimization* 4359: 187–199.
- Laporte, G., Á. Marín, J. A. Mesa, and F. Perea. 2011. "Designing Robust Rapid Transit Networks with Alternative Routes." *Journal of Advanced Transportation* 45: 54–65.
- Laporte, G., and J. A. Mesa. 2014. "The Design of Rapid Transit Networks." Les Cahiers du Gerard G-2014-08. <https://www.gerad.ca/fr/papers/G-2014-08/view>.
- Leiva, C., J. C. Muñoz, R. Giesen, and H. Larrain. 2010. "Design of Limited-Stop Services for an Urban Bus Corridor with Capacity Constraints." *Transportation Research Part B* 44: 1186–1201.
- Li, Z.-C., Z.-K. Li, and W. H. K. Lam. 2014. "An Integrated Design of Sustainable Land Use and Transportation System with Uncertainty in Future Population." *Transportmetrica A: Transport Science* 10 (2): 160–185.

- Liebchen, C., M. Lübbecke, R. Möhring, and S. Stiller. 2009. "The Concept of Recoverable Robustness, Linear Programming Recovery and Railway Applications." In *Robust and Online Large-Scale Optimization*, Volume 5868 of Lecture Notes in Computer Science, edited by R. Ahuja, R. Möhring, and C. Zaroliagis, 1–27. Berlin: Springer.
- Marín, Á. 2007. "An Extension to Rapid Transit Network Design Problem." *Top* 15: 231–241.
- Marín, Á., and P. Jaramillo. 2008. "Urban Rapid Transit Network Capacity Expansion." *European Journal of Operational Research* 191 (1): 45–60.
- Marín, Á., and R. García-Ródenas. 2009. "Location of Infrastructure in Urban Railway Networks." *Computers & Operations Research* 36: 1461–1477.
- Mauttone, A., and M. E. Urquhart. 2009. "A Multi-Objective Metaheuristic Approach for the Transit Network Design Problem." *Public Transport* 1 (4): 253–273.
- Nielsen, L. K., L. Kroon, and G. Maróti. 2012. "A Rolling Horizon Approach for Disruption Management of Railway Rolling Stock." *European Journal of Operational Research* 220 (2): 496–509.
- Michaelis, M., and A. Schöbel. 2009. "Integrating Line Planning, Timetabling, and Vehicle Scheduling: A Customer-Oriented Heuristic." *Public Transport* 1 (3): 211–232.
- Szeto, W. Y., and H. K. Lo. 2005. "Strategies for Road Network Design Over Time: Robustness Under Uncertainty." *Transportmetrica* 1 (1): 47–63.
- Szeto, W. Y., and A. B. Wang. 2015. "Price of Anarchy for Reliability-Based Traffic Assignment and Network Design." *Transportmetrica A: Transport Science* 11 (7): 603–635.
- Vuchic, V. R. 2005. *Urban Transit: Operations, Planning, and Economics*. Hoboken: Wiley.
- Walteros, J. L., A. L. Medaglia, and G. Riaño. 2013. "Hybrid Algorithm for Route Design on Bus Rapid Transit Systems." *Transportation Science* 49 (1): 66–84. doi:10.1287/trsc.2013.047.
- Wei, F., and R. B. Machemehl. 2004. "Optimal Transit Route Network Design Problem: Algorithms, Implementations and Numerical Results." Research SWUTC/04/167244-1, University of Texas.